

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-268922

(43)Date of publication of application : 20.09.2002

(51)Int.Cl.

G06F 11/34
G06F 13/00

(21)Application number : 2001-066617

(71)Applicant : NTT DATA CORP

(22)Date of filing : 09.03.2001

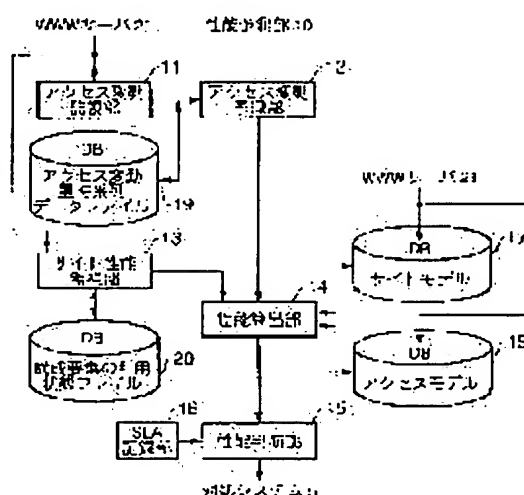
(72)Inventor : SAKATA YUJI
KOYAMA TAKAO

(54) PERFORMANCE MONITORING DEVICE OF WWW SITE

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a performance monitoring device of a WWW site to realize a preventative countermeasure for maintenance of performance.

SOLUTION: The performance monitoring device 1 of the WWW site is constituted of an access fluctuation monitoring part 11 to fetch and store access quantity in the WWW site per unit time as time-series data of access fluctuation, an access fluctuation predicting part 12 to predict the access fluctuation in the future by statistical technique based on the stored pieces of time-series data, a site performance monitoring part 13 to acquire and store performance data by which use states of components in the WWW site from the components, a performance calculating part 14 to calculate performance values by every access based on a model in which the components of the site are abstracted from prediction data of the access fluctuation and performance monitoring data of the site, and a performance judging part 15 to judge whether or not prescribed performance is maintained based on the performance values of the calculated performance values by every access.



LEGAL STATUS

[Date of request for examination]

14.06.2005

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

【特許請求の範囲】

【請求項 1】 単位時間当たりのアクセス変動の予測データと、WWWサイトにおける構成要素の利用状況から成る監視データから、前記WWWサイトの構成要素を抽象化したモデルを元に、定められた性能を維持できるかを判断する性能予測手段を備えたことを特徴とするWWWサイトの性能監視装置。

【請求項 2】 前記性能予測手段で定められた性能を維持できないと判断されたときに、前記性能を維持できない要因となる前記WWWサイトの構成要素を推測することを特徴とする請求項 1 に記載のWWWサイトの性能監視装置。

【請求項 3】 前記性能予測手段は、WWWサイトにおける単位時間当たりのアクセス量をアクセス変動の時系列データとして取り込み蓄積するアクセス変動監視手段と、前記蓄積された時系列データを元に統計的な手法で将来的なアクセス変動を予測するアクセス変動予測手段と、前記WWWサイトにおける構成要素の利用状態を把握する性能データを前記構成要素から取得して蓄積するサイト性能監視手段と、前記アクセス変動の予測データとサイトの性能監視データとから、サイトの構成要素を抽象化したモデルに基づきアクセス毎の性能値を算出する性能算出手段と、前記性能算出手段で算出されたアクセス毎の性能値を元に定められた性能が維持できるかを判断し、かつ、前記性能を維持できない要因となる前記WWWサイトの構成要素を推測する性能判断手段と、を備えたことを特徴とする請求項 1 または 2 に記載のWWWサイトの性能監視装置。

【請求項 4】 前記アクセス変動予測手段は、前記アクセス変動監視手段が蓄積した時系列データの回帰分析により求めた予測モデルから指定時間後のアクセス変動を求め、前記回帰分析時に複数の回帰モデルを同時に計算し、前記アクセス変動結果との相関関数が最も高い回帰モデルを予測モデルとして使用することを特徴とする請求項 3 に記載のWWWサイトの性能監視装置。

【請求項 5】 前記サイトの構成要素を抽象化したモデルは、サイトの構成要素を待ち行列モデルでモデル化したサイトモデルと、前記サイトへのアクセス毎、前記サイトの構成要素への負荷を見積もったアクセスモデルから成り、前記性能算出手段は、前記アクセス変動予測とサイトの性能監視データとから、求めたい時点におけるアクセスレスポンスタイムを前記各モデルを元に算出することを特徴とする請求項 3 に記載のWWWサイトの性能監視装置。

【請求項 6】 前記アクセスレスポンスタイムは、前記アクセス変動予測に基づき、全てのアクセスをその生起時刻が記述された事象カレンダーとして作成し、前記指定の時間間隔において状態が定常的であると仮定してシス

テム内部を任意に区切り、その内部で生じる事象に対して如何なる順序で逐次処理を行なっても問題なしと仮定して生じる事象を、前記指定の時間を 1 ステップとして前記求めたい時間までシミュレーションのための計算を行なうことにより求めることを特徴とする請求項 5 に記載のWWWサイトの性能監視装置。

【発明の詳細な説明】**【0001】**

【発明の属する技術分野】 この発明は、WWW (World Wide Web) サイトの性能の変動を事前に予測し、その結果から予防的な対処を行うことで、WWWサイトの定められた性能を維持するWWWサイトの性能監視装置に関する。

【0002】

【従来の技術】 現在、WWWを介したビジネスが急激な勢いで立ち上がろうとしている。このようなWWWサイトにおいては、顧客の確保、維持のため、高性能で信頼性のあるサービスを提供することは不可欠と考えることができる。そのため、現在、顧客に対して、そのWWWサイトのサービス品質を定量的に設定し、それを遵守するような S L M (Service Level Management) と言われる運用管理手法が注目されている。このような定められた一定品質のサービスを維持するためには、なんらかの障害に対して、事後的な対応を行うのではなく、問題が発生する前に予防的な対処を行う必要がある。

【0003】

【発明が解決しようとする課題】 上記したように、商用のWWWサイトにおいて、性能維持管理は非常に重要である。しかしながら最近のWWWサイトは、ネットワーク、サーバ、データベース、アプリケーション等様々な要素から構成されるものであり、利用者が体感する性能を全体として維持していくことは非常に困難である。WWWサイトの性能維持管理のためには、性能の計測と性能悪化時の対処が必須であるが、更に性能悪化を予測し、迅速に予防的な対処を行なうことができると一定の性能品質を維持することが可能になる。このような性能維持管理のために既存の管理製品には、従来、以下に

(1)、(2)として示す手法を用いてその管理を行なうものがあつた。

(1) 定められたWWWサイトへのアクセスのレスポンスタイムを一定間隔で計測、表示し、決められた値を越えると警告を発する。また、その計測データを時系列データとして蓄積し、何らかの統計的解析手法で将来の性能変動を予測する。

(2) ネットワーク、サーバ、アプリケーション等、WWWサイトの構成要素の、ネットワーク利用率、CPU 使用率等内部状態値を監視し、それらが定められた閾値を越えると警告を発する。また、その計測データを時系列データとして蓄積して何らかの統計的解析手法により将来の性能変動を予測する。

【0004】しかしながら、上記（１）に示す手法によれば、WWWサイトの外部からのレスポンスタイムを計測するために、何らかの性能低下が見られた場合、それがWWWサイトを構成するどの要素に起因する問題なのか特定が困難である。また、システムに何らかの障害が発生した場合を除けば、一般的に性能品質の低下は、ユーザ動向の変化に起因するものと考えることができる。レスポンスタイムは、WWWサイト内の複雑な構成要素による処理の結果定まるものであり、レスポンスタイム自身の時系列変化にモデルを仮定することは難しい。従って、レスポンスタイム値のみを用いて統計的な手法により、レスポンスタイム値の変化を予測するための方法は一般化が困難であるといえる。一方、上記した（２）に示す手法によれば、WWWサイトは、各構成要素が複雑に関連しているため、構成要素の個々の状態を監視しているだけでは、WWWサイト全体としての性能品質を把握することは難しい。また、同様の理由で内部状態のみを用いてWWWサイト船体の性能品質を予測することは困難であるといえる。

【0005】本発明は上記諸々の事情に鑑みてなされたものであり、WWWサイトを、その構成要素（リソース）を考慮できる形態でモデル化し、WWWサイトへのアクセス変動からアクセスのレスポンスタイム等の性能データを算出するための手段を実装し、これを用いてWWWサイトのアクセス変動を予測することによりWWWサイトの性能変動を予測し、性能維持のための予防的対処を実現するWWWサイトの性能監視装置を提供することを目的とする。また、上記の結果から定められた性能を将来的に維持できるか否かを判断し、維持できないと判断されたときにWWWサイトを構成するどの要素が性能悪化に影響を与えているかを上記性能データに基づき判断して外部システムに通知するWWWサイトの性能監視装置を提供することも目的とする。

【0006】

【課題を解決するための手段】上記した課題を解決するために本発明は、単位時間当たりのアクセス変動の予測データと、WWWサイトにおける構成要素の利用状況から成る監視データから、前記WWWサイトの構成要素を抽象化したモデルを元に、定められた性能を維持できるか否かを判断する性能予測手段を備えたことを特徴とする。

【0007】また、本発明のWWWサイトの性能監視装置において、前記性能予測手段で定められた性能を維持できないと判断されたときに、前記性能を維持できない要因となる前記WWWサイトの構成要素を予測することを特徴とする。

【0008】また、本発明のWWWサイトの性能監視装置において、前記性能予測手段は、WWWサイトにおける単位時間当たりのアクセス量をアクセス変動の時系列データとして取り込み蓄積するアクセス変動監視手段

と、前記蓄積された時系列データを元に統計的な手法で将来的なアクセス変動を予測するアクセス変動予測手段と、前記WWWサイトにおける構成要素の利用状態を把握する性能データを前記構成要素から取得して蓄積するサイト性能監視手段と、前記アクセス変動の予測データとサイトの性能監視データとから、サイトの構成要素を抽象化したモデルに基づきアクセス毎の性能値を算出する性能算出手段と、前記性能算出手段で算出されたアクセス毎の性能値を元に定められた性能が維持できるか否かを判断し、かつ、前記性能を維持できない要因となる前記WWWサイトの構成要素を推測する性能判断手段と、を備えたことを特徴とする。

【0009】また、本発明のWWWサイトの性能監視装置において、前記アクセス変動予測手段は、前記アクセス変動監視手段が蓄積した時系列データの回帰分析により求めた予測モデルから指定時間後のアクセス変動を求め、前記回帰分析時に複数の回帰モデルを同時に計算し、前記アクセス変動結果との相関関数が最も高い回帰モデルを予測モデルとして使用することを特徴とする。

【0010】また、本発明のWWWサイトの性能監視装置において、前記サイトの構成要素を抽象化したモデルは、サイトの構成要素を待ち行列モデルでモデル化したサイトモデルと、前記サイトへのアクセス毎、前記サイトの構成要素への負荷を見積もったアクセスモデルから成り、前記性能算出手段は、前記アクセス変動予測とサイトの性能監視データとから、求めたい時点におけるアクセスレスポンスタイムを前記各モデルを元に算出することを特徴とする。

【0011】また、本発明のWWWサイトの性能監視装置において、前記アクセスレスポンスタイムは、前記アクセス変動予測に基づき、全てのアクセスをその生起時刻が記述された事象カレンダーとして作成し、前記指定の時間間隔において状態が定常的であると仮定してシステム内部を任意に区切り、その内部で生じる事象に対して如何なる順序で逐次処理を行なっても問題なしと仮定して生じる事象を、前記指定の時間を１ステップとして前記求めたい時間までシミュレーションのための計算を行なうことにより求めることを特徴とする。

【0012】上記構成により、単位時間当たりのアクセス変動の予測データと、WWWサイトにおける構成要素の利用状況から成る監視データからサイトの構成要素を抽象化したモデルに基づきアクセス毎の性能値を算出し、算出されたアクセス毎の性能値を元に定められた性能が維持できるか否かを判断すると共に、その性能を維持できない要因となるWWWサイトの構成要素を推測する。このことにより、WWWサイトの性能悪化を予測して予防的な対処を行なうことができ、アクセス変動から予測することによってより精度の高い予測を可能とし、また、WWWサイトの構成要素をモデル化してそのモデルから計算しているため、何らかの性能悪化が予想され

た場合、構成要素のどの部位がボトルネックになっているかを推測することが容易に実現できる。

【0013】

【発明の実施の形態】以下、図面を参照して本発明の実施の形態について説明する。まず、図1により本実施の形態の性能監視装置が用いられるシステム全体の構成例を説明する。図1において、符号1は本実施の形態の性能監視装置である。符号2は、性能監視対象のWWWサイトであって、アクセスするクライアントに対して情報を開示する情報提供システムである。符号3は、クライアント（WWWサイト2の顧客）である。符号4は、WWWサイト2とクライアント3を接続するインターネット（コンピュータネットワーク）である。符号5は、対処システムである。対処システムは、性能が悪化したときに、性能維持のために動的に何らかの操作ができるシステムである。

【0014】WWWサイト2は、WWWサーバ21と、コンテンツデータベース22と、ネットワーク機器23とから構成されている。WWWサーバ21は、クライアント3からの情報開示要求を受け付け、該当する情報をクライアント3に対して提供する。コンテンツデータベース22は、WWWサイト2が提供する情報コンテンツを予め記録するデータベースである。ネットワーク機器23は、インターネット4とWWWサーバ21を接続するルータを含むコンピュータネットワーク用のインタフェース機器である。

【0015】図1では、クライアント3がインターネット4を介してWWWサイト2へアクセスし、情報の開示を要求すると、WWWサイト2は、コンテンツデータベース22の中から要求された情報を抽出し、WWWサーバ21により提供する。性能監視装置1は、単位時間当たりのアクセス変動の予測データと、WWWサーバ21における構成要素の利用状況から成る監視データから、WWWサーバ21の構成要素を抽象化したモデルを元に、定められた性能を維持できるか否かを判断する。また、定められた性能を維持できないと判断されたときに、性能を維持できない要因となる前記WWWサーバ21の構成要素を推測し、対処システム5へ供給する。対処システム5は、動的に、サーバ等のリソースを増加、あるいは減少させることができ、あるいは、WWWサイト2へのアクセスを選別、制限できる仕組みを備えるものとするが、本発明要旨には直接関係しないため、ここでの説明は要しない。

【0016】図2は、図1に示す性能監視装置1の内部構成を機能展開して示した図である。以下に示す各ブロックは、具体的には、CPUならびにメモリを含む周辺LSIで構成され、CPUがメモリに記録されたプログラムを読み出し、逐次実行することによってその機能を実現するものである。ここでは、性能監視装置1が持つコンポーネントのうち、のうち性能予測部10のみが抽

出して示されている。性能予測部10は、アクセス変動監視部11、アクセス変動予測部12、サイト性能監視部13、性能算出部14、性能判断部15、SLA記録部16、サイトモデルDB17、アクセスモデルDB18、アクセス変動量時系列データファイル19（以下、単にDB19という）、構成要素の利用状態ファイル20（以下、単にDB20という）で構成される。

【0017】アクセス変動監視部11は、WWWサイト2における単位時間当たりのアクセス量をアクセス変動の時系列データとして取り込み、DB19に蓄積する機能を有する。アクセス変動予測部12は、DB19に蓄積された時系列データを元に統計的な手法で将来的なアクセス変動を予測する機能を有する。サイト性能監視部13は、WWWサイト2における構成要素の利用状態を把握する性能データを構成要素から取得してDB20に蓄積する機能を有する。また、アクセス変動予測部12は、後述するように、アクセス変動監視部11が蓄積した時系列データの回帰分析により求めた予測モデルから指定時間後のアクセス変動を求め、回帰分析時に複数の回帰モデルを同時に計算し、アクセス変動結果との相関関数が最も高い回帰モデルを予測モデルとして使用する。

【0018】性能算出部14は、アクセス変動の予測データとサイトの性能監視データとから、サイトの構成要素を抽象化したモデルに基づきアクセス毎の性能値を算出する機能を有する。なお、サイトの構成要素を抽象化したモデルは、サイトの構成要素を待ち行列モデルでモデル化したサイトモデル（サイトモデルDB17）と、サイトへのアクセス毎、サイトの構成要素への負荷を見積もったアクセスモデル（アクセスモデルDB18）から成り、性能算出部14は、アクセス変動予測とサイトの性能監視データとから、求めたい時点におけるアクセスレスポンスタイムを各モデルを元に算出する。更に、アクセスレスポンスタイムは、アクセス変動予測に基づき、全てのアクセスをその生起時刻が記述された事象カレンダーとして作成し、指定の時間間隔において状態が定常的であると仮定してシステム内部を任意に区切り、その内部で生じる事象に対して如何なる順序で逐次処理を行なっても問題なしと仮定して生じる事象を、先に指定の時間を1ステップとして求めたい時間までシミュレーションのための計算を行なうことによって求める。

【0019】性能判断部15は、性能算出部14で算出されたアクセス毎の性能値を元に、SLA記録部16に設定されてある、定められた性能が維持できるか否かを判断し、かつ、性能を維持できない要因となるWWWサイト2の構成要素を推測する機能を有する。なお、SLA記録部16には、システムの可用性、性能等の目標量を定量的に取り決めたものである。

【0020】図3は、本発明の性能監視装置の基本動作を説明するために引用したフローチャートである。以

下、図 3 に示すフローチャートを参照しながら本発明の性能監視装置 1 の基本動作について説明する。まず、アクセス変動監視部 11 は、定時間間隔でアクセス、もしくはグループ化されたアクセス種毎の単位時間当たりのアクセス数を WWW サーバ 21 から取得する（ステップ S31）。そして、アクセス変動の時系列データとして DB19 に蓄積する（ステップ S32）。次に、アクセス変動予測部 12 では、DB19 に蓄積されたアクセス変動データを元に将来的なアクセス変動を後述する回帰分析等の統計的手法を用いて予測する（ステップ S33）。一方、サイト性能監視部 13 では、サイトを構成するリソース（構成要素）の現在の利用状態を把握できるような性能データに関し、それらリソースから定期的に取得して DB20 に蓄積する（ステップ S34）。

【0021】性能算出部 14 では、アクセス変動予測部 12 から得られるアクセス変動の予測データと、サイト性能監視部 13 から得られるサイトの性能監視データから、サイトのリソース構成を抽象化したモデル（サイトモデル DB17、アクセスモデル DB18）を元にアクセス毎のレスポンスタイム値を算出する（ステップ S35）。なお、サイトを抽象化したモデルは、WWW サーバ 21、コンテンツデータベース 22 等、サイトの構成要素を待ち行列モデルでモデル化したサイトモデルと、そのサイトへのアクセスがこのモデルの個々の要素にどの程度負荷をかけるかを示したアクセスモデルから構成され、それぞれ、サイトモデル DB17、アクセスモデル DB18 に格納される。これらモデルは、事前の負荷試験、個々のリソースの基礎性能データ、プログラムロジック等から本発明装置を稼動させる前に事前に構築されるものとする。レスポンスタイム値の算出のためにシミュレーション、統計的手法、解析的手法等を用いることができる。

【0022】次に、性能判断部 15 では、性能算出部 14 で算出されたアクセス毎の性能予測結果を元に、SLA 記録部 16 に記録された、事前に定められた性能品質を維持できるか否かを判断する（ステップ S36）。ここで維持できると判断された場合は、ステップ S31 以降の処理を繰り返し、維持できないと判断された場合は、管理者もしくは、外部接続された対処システム 5 に通知する。そして、その際に、サイトモデルの要素のどの部分で性能が悪化しているかを示す（ステップ S37）ことにより、性能悪化の根本原因の切り分けも示すことができる。

【0023】図 4～図 14 は、本発明実施形態の動作を説明するために引用した図であり、それぞれ、性能モデルを概念的に示した図（図 4）、回帰分析によるアクセス変動の予測を説明するために引用した図（図 5、図 6、図 7）、待ち行列の定義を説明するために引用した概念図（図 8）、待ち行列網のモデル化を説明するために引用した図（図 9）、WWW サイトのモデル化の一例

ならびにそのときの WWW サイトで利用できるアクセスの種類を説明するために引用した図（図 10）、モデル化に用いられる設定値の例（図 11）、事象カレンダー（図 12）、時間追跡法を用いたシミュレーションによるレスポンスタイム値の計算処理の流れを示すフローチャート（図 13）、性能判断部 15 によるボトルネック部の切り離しを説明するために引用した図（図 14）である。

【0024】以下、図 4～図 14 を参照しながら図 1～図 3 に示す本発明実施形態の動作について詳細に説明する。本発明の基本思想は、例えば、回帰分析によるアクセス変動予測を行い、アクセス変動予測値と内部状態値からレスポンスタイムを予測できるような性能モデルを構築し、この性能モデルは、WWW サイト 2 の内部リソースをモデル化したものとするることにより、ボトルネック部を把握できるようにしたことにある。図 4 にその性能モデルが模式化して示されている。

【0025】次に、図 5～図 8 を用いてアクセス変動予測部 12 の回帰分析によるアクセス変動の予測について説明する。図 5 は、アクセス変動予測部 12 の動作を説明するフローチャート、図 6 は、アクセス変動予測部 12 の予測モデルの導出と予測例を説明する模式図、図 7 は、アクセス変動予測部 12 における予測モデルの導出と選択のための情報テーブルの一例を説明する模式図である。アクセス変動監視部 11 では、まず、WWW サーバ 21 のアクセスログから単位時間当たりのアクセス数の時系列データを取得し、DB19 に蓄積する。アクセス変動予測部 12 は、DB19 からその時系列データにおける一定の履歴範囲のデータを取り出し（図 5 のステップ S331）、複数モデル関数に回帰分析によって当てはめることで最適なモデル関数を導出する。例えば、このモデルの確定は、以下のモデルの確定方法により計算する。

（1）ある一定の範囲で、測定時間とアクセス数の値が線形の関係にある。このモデルの場合、両者の関係は数式

【数 1】

$$y = ax + b$$

ここで y はアクセス数、 x は予測範囲時間という形式で表現される。

（2）ある一定の範囲で、測定時間とアクセス数の値が二次曲線の関係にある。このモデルの場合、両者の関係は数式

【数 2】

$$y = ax^2 + b$$

ここで y はアクセス数、 x は予測範囲時間という形式で表現される。

【0026】これらのモデルの係数を、与えられたデー

タから回帰分析を行って導出し、モデルを確定させる(ステップS333)。また、各モデルは上記のような単純な回帰モデルだけではなく、指数関数や対数関数等で表される非線形回帰モデルの他、ニューラルネットワークなどを用いる方法が考えられる。そして、各モデルによって計算される予測モデル理論値とアクセス実績値とを比較し、最も適合するモデルを今回の予測モデルとして決定する(ステップS334)。ここで、最も適合するモデルを決定する方法には、例えば予測モデル理論値とアクセス実績値との相関係数を求め、相関係数が最も大きいモデルを採用する方法などがある。図6に示す例では、二次曲線モデル(二乗モデル)が最も適合するモデルとして特定される。

【0027】本実施の形態では、図7で示した<線形モデル>、<二乗モデル>、<ルートモデル>の三種類のモデルで定義される関数で表現できると考える。そして、 x 、 y の組み合わせが最もモデルに適合するような最適な係数 a 、 b を、例えば、最小近似法により求める。それぞれのモデルの相関係数が求まり、ここでは、相関係数の値が最も高いルートモデルが今回最も適合する予測モデルとして決定される。上記のように決定された予測モデルを用い、アクセス変動予測部12は、設定された予測範囲におけるアクセス変動の予測値を計算する(ステップS335)。

【0028】次に、性能算出部14におけるレスポンスタイム算出の際に用いられる待ち行列の定義について説明する。ここでは、アクセス、サービス機構、サービス窓口等の術語を使って待ち行列が説明されている。待ち行列モデルを規定するためには、母集団の大きさ、アクセスの到着、サービス時間分布、システム最大容量、窓口の数、待ち行列の基準を考える必要がある。待ち行列理論には数多くのモデルが存在するが、ここでは、説明を簡略化する意味で実際の待ち行列理論において利用されない用語も含まれていることを断っておく。すなわち、待ち行列理論において定義される待ち行列を正しく定義したものではない。

【0029】以下、待ち行列へ到着するアクセスの動作について、図8を参照しながら説明する。まず系内にある待ち行列に空きがある場合、アクセスが待ち行列に到着する(ステップS81)。サービス窓口が全て塞がっている場合、アクセスは、待ち行列内に到着順に並ぶ(ステップS82)。そして、サービス窓口が空いた場合、待ち行列の先頭のアクセスが空いたサービス窓口に入る(ステップS83)。サービス窓口で、アクセスは、サービス機構の性能と同時アクセス数から導出されるサービス窓口の処理能力に応じたサービス量を受ける(ステップS84)。そして、アクセスがアクセス毎に定められた負荷を(サービス量)をサービス窓口で受けるとそのアクセスは、サービス機構から退出する(ステップS85)。なお、ステップS81～S85までに要

した時間を系内時間、ステップS81～S83までに要した時間を待ち時間、ステップS83～S85までに要した時間をサービス時間という。

【0030】次に、上記した待ち行列の網のモデル化について説明する。ここでは、待ち行列の設定条件によって図9の(a)、(b)、(c)のように分類する。図9(a)、(b)、(c)は、ともに待ち行列網の構成と、縦軸に処理能力、横軸に同時アクセス数を目盛ったグラフが示されている。すなわち、(a)は、待ち行列がなく一定の待ち時間ありのカテゴリ、(b)は、待ち行列があり一定の待ち時間があるカテゴリ、(c)は、待ち行列があり、共有リソースを用いることから同時サービス数に反比例した処理能力を持つカテゴリに区分される。従って、(b)において、サービス窓口の処理能力は性能と等しく、(c)において、サービス窓口の処理能力は、性能/同時アクセス数で示される。

【0031】図10に、WWWサイトのモデル化例が示されている。図において、(a)は、WWWサイト2のシステム構成を、(b)は、WWWサイト2で利用できるアクセスの種類(IDとURL(Uniform Resource Locator)が示されている)を、(c)は、WWWサイト2のシステム構成要素に対応して作成された待ち行列網モデルである。図中、図1に示すブロックと同一番号の付されたブロックは図1に示すそれと同じとする。ここでは、WWWサイト2は、コンテンツデータベース(DB)22ならびにWWWサーバ21およびアプリケーション(AP)サーバ100が主要素として示され、両者は、100Mbpsの容量を持つLAN(Local Area Network)によって共通接続されている。

【0032】図11に、モデル化に用いられた設定値のそれぞれが示されている。図11において、(a)は、WWWサイト2のプログラムロジックと負荷試験から事前に求められるURL毎の設定値であり、図10に示すシステムに適用される。ここでは、WWWサーバ21/APサーバ100と、コンテンツデータベース22に区分して示してある。(b)は、サービス機構としてのWWW21/APサーバ100、あるいはコンテンツデータベース22の性能値である。(c)は、待ち行列の設定値であり、WWWサーバ21/APサーバ100、コンテンツデータベース22、クライアント3毎の最大待ち行列数、同時処理サービス数、共有リソース、待ち時間が設定される。なお、ここで、「性能値」とは、待ち行列で示されたサーバの処理能力を示し、「設定値」とは、アクセスの種類によって決定されるアクセスの待ち行列網内での振る舞いと個々の待ち行列網にかかる負荷を指す。

【0033】次に、性能算出部14によるレスポンスタイム値の計算の仕方について説明する。性能算出部14は、アクセス変動の予測データとサイトの性能監視データとから、サイトの構成要素を抽象化したモデルに基づ

きアクセス毎の性能値を算出する機能を有することは上記したとおりである。ここでは、時間追跡法を用いたシミュレーションによりレスポンスタイムを計算するものとする。全てのアクセスは、その生起時刻が記述された事象カレンダーとして記載される。なお、事象カレンダーは、アクセス変動予測部 12 により予測されたアクセス変動値から作成されるものとする。図 12 にその事象カレンダー 120 の一例が示されている。ここでは、ある短い時間 Δt に注目して、その間に状態が定常的であると仮定し、また、システム内部を幾つかに区切ってその内部で生じる事象に対してどのような順序で逐次処理を行なっても問題ないと仮定して生じる事象を計算するものとした。このような条件の下、 Δt を 1 ステップとして求めたい時間 T_{END} まで計算を行なう処理が図 13 に示すフローチャートに記述されている。

【0034】図 13 に示すフローチャートにおいて、まず、変数 t にアクセスの生起時刻をセットする（ステップ S351）。そして、待ち行列に基づき、サービス窓口におけるサービス処理を行なうと共に、待ち行列に存在するアクセスのサービス窓口への割り当てを行なう（ステップ S353、S353）。次に、事象カレンダー 120 からのアクセスを取得し（ステップ S354）、再度待ち行列に対するアクセスのサービス窓口への割り当てを行なう（ステップ S355）。そして、先の変数 t を Δt 分だけ更新してステップを進め（ステップ S356）、以上の処理を求めたい時間 T_{END} まで繰り返し計算を行なう（ステップ S357）。

【0035】最後に、性能判断部 15 によるボトルネックの切り分け例についてについて説明する。性能判断部 15 は、性能算出部 14 で算出されたアクセス毎の性能値を元に、SLA 記録部 16 に設定されてある、定められた性能が維持できるか否かを判断し、かつ、性能を維持できない要因となる WWW サイト 2 の構成要素を推測する。WWW サイト 2 におけるレスポンスタイム値は、WWW サイト 2 内のそれぞれの待ち行列における系内時間の合計として表現することができる。従って、個々の待ち行列の系内時間から WWW サイト 2 内のどの部分で処理に時間を要しているかを把握することができる。図 14 に ID 毎の待ち行列における系内時間と、レスポンスタイム値が示されているが、ここでは、コンテンツ DB 22 の処理時間がレスポンスタイムに影響していることが理解できる。

【0036】なお、図 2 に示す性能監視装置 1 は、その機能を実現するためのプログラムを、コンピュータ読みとり可能な記録媒体に記録して、この記録媒体に記録されたプログラムをコンピュータシステムに読み込ませ、実行することにより、上述の各装置における機能を実現しても良い。

【0037】ここで、上記「コンピュータシステム」とは、OS や周辺機器等のハードウェアを含み、さらに W

WW (World Wide Web) システムを利用している場合であれば、ホームページ提供環境（あるいは表示環境）も含むものとする。また、「コンピュータ読みとり可能な記録媒体」とは、フロッピー（登録商標）ディスク、光磁気ディスク、OM、CD-ROM 等の可搬媒体、コンピュータシステムに内蔵されるハードディスク等の記憶装置のことをいう。更に、「コンピュータ読みとり可能な記録媒体」とは、インターネット等のコンピュータネットワークや電話回線等の通信回線を介してプログラムを送信する場合のように、短時間の間、動的にプログラムを保持するもの（伝送媒体もしくは伝送波）、その場合のサーバやクライアントとなるコンピュータシステム内部の揮発性メモリのように、一定時間プログラムを保持しているものも含むものとする。

【0038】また、上記プログラムは、前述した機能の一部を実現するためのものであっても良く、更に前述した機能をコンピュータシステムに既に記憶されているプログラムとの組み合わせで実現できるもの、いわゆる差分ファイル（差分プログラム）であっても良い。

【0039】

【発明の効果】以上説明のように本発明は、WWW サイトの内部構成モデルを保持し、また、定期的にサイトアクセスの変動を予測することによってサイトの性能を予測し、また、定められた性能を維持できないと判断された場合に WWW サイト内部のボトルネック部分に対処システムに通知してその改善を促がすものであり、このために、単位時間当たりのアクセス変動の予測データと、WWW サイトにおける構成要素の利用状況から成る監視データからサイトの構成要素を抽象化したモデルに基づきアクセス毎の性能値を算出し、算出されたアクセス毎の性能値を元に定められた性能が維持できるか否かを判断すると共に、その性能を維持できない要因となる WWW サイトの構成要素を推測するものとした。

【0040】このことにより、WWW サイトの性能悪化を予測して予防的な対処を行なうことができる。本発明によれば、アクセス変動から予測することによってより精度の高い予測を可能とし、また、WWW サイトの構成要素をモデル化してそのモデルから計算しているため、何らかの性能悪化が予想された場合、構成要素のどの部位がボトルネックになっているかを推測することが可能となる。従って、WWW サイトの性能の変動を事前に予測し、その結果から予防的な対処を行うことで、WWW サイトの定められた性能を維持することができるという効果が得られる。

【図面の簡単な説明】

【図 1】 本発明実施形態の性能監視装置が用いられるシステム全体の構成例を説明するために引用した図である。

【図 2】 図 1 に示す性能監視装置の内部構成を機能展開して示したブロック図である。

【図 3】 本発明実施形態の基本動作を説明するために引用したフローチャートである。

【図 4】 本発明実施形態の動作を説明するために引用した図であり、性能モデルを概念的に示した図である。

【図 5】 本発明実施形態の動作を説明するために引用した図であり、回帰分析によるアクセス変動の予測を説明するために引用した図である。

【図 6】 本発明実施形態の動作を説明するために引用した図であり、回帰分析によるアクセス変動の予測を説明するために引用した図である。

【図 7】 本発明実施形態の動作を説明するために引用した図であり、回帰分析によるアクセス変動の予測を説明するために引用した図である。

【図 8】 本発明実施形態の動作を説明するために引用した図であり、待ち行列の定義を説明するために引用した概念図である。

【図 9】 本発明実施形態の動作を説明するために引用した図であり、待ち行列網のモデル化を説明するために引用した図である。

【図 10】 本発明実施形態の動作を説明するために引用した図であり、WWWサイトのモデル化の一例、ならびにそのときのWWWサイトで利用できるアクセスの種類を説明するために引用した図である。

【図 11】 本発明実施形態の動作を説明するために引用した図であり、モデル化に用いられる設定値の例を示す図である。

【図 12】 本発明実施形態の動作を説明するために引用した図であり、事象カレンダーの構成例を示す図であ

る。

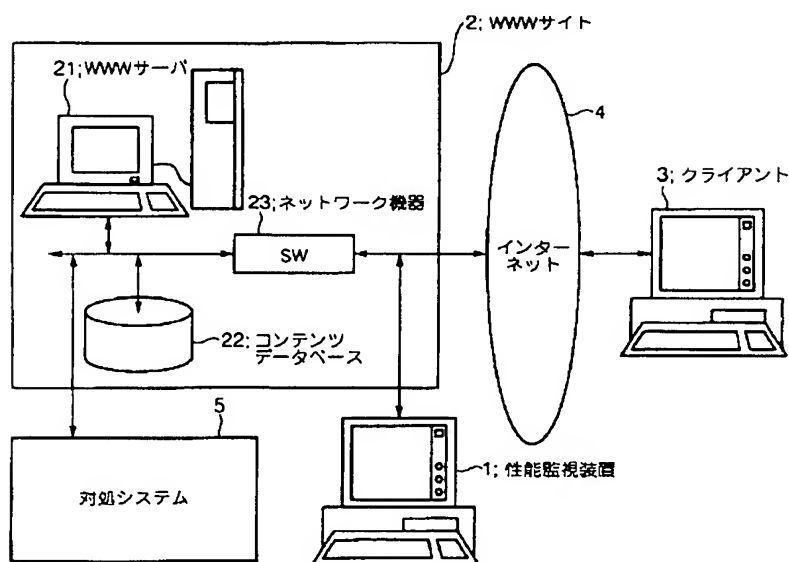
【図 13】 本発明実施形態の動作を説明するために引用した図であり、時間追跡法を用いたシミュレーションによるレスポンスタイム値の計算処理の流れを示すフローチャートである。

【図 14】 本発明実施形態の動作を説明するために引用した図であり、性能判断部によるボトルネック部の切り離しを説明するために引用した図である。

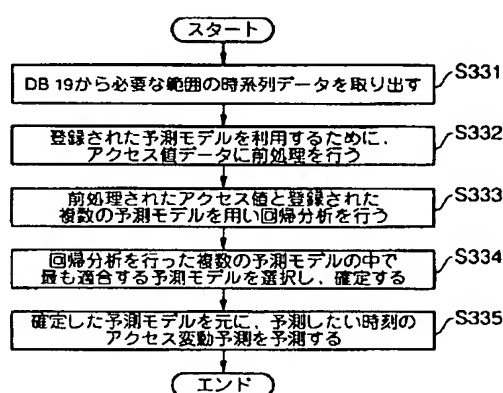
【符号の説明】

- 1 性能監視装置
- 2 WWWサイト
- 3 クライアント
- 4 インターネット
- 5 対処システム
- 11 アクセス変動監視部
- 12 アクセス変動予測部
- 13 サイト性能監視部
- 14 性能算出部
- 15 性能判断部
- 16 SLA記録部
- 17 サイトモデルDB
- 18 アクセスモデルDB
- 19 アクセス変動量時系列データファイル
- 20 構成要素の利用状態ファイル
- 21 WWWサーバ
- 22 コンテンツデータベース
- 23 ネットワーク機器
- 24 ネットワーク機器

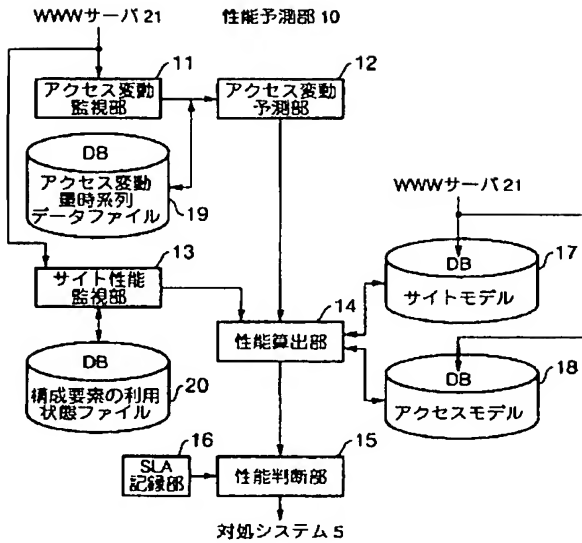
【図 1】



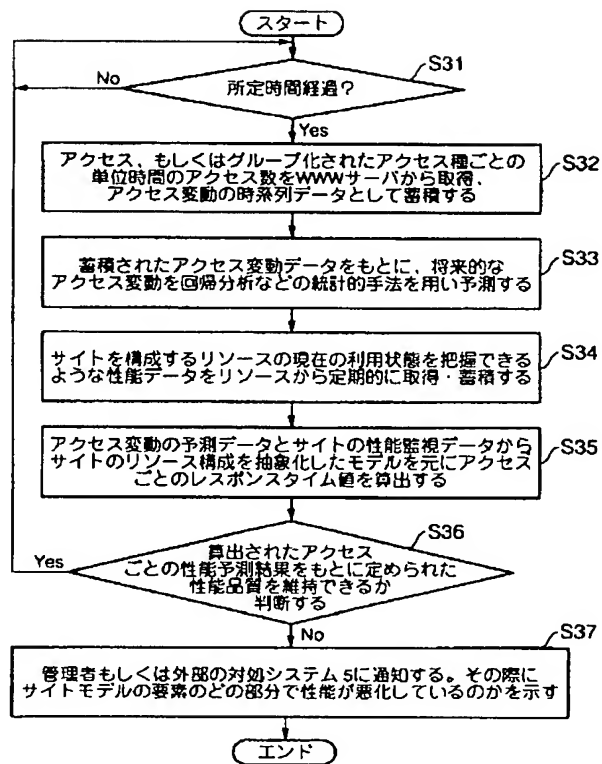
【図 5】



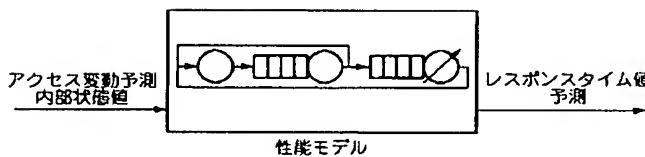
【図2】



【図3】



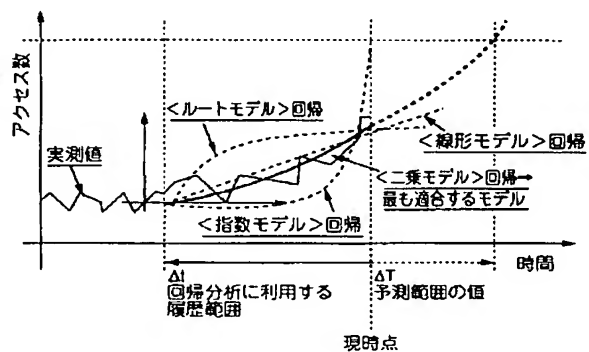
【図4】



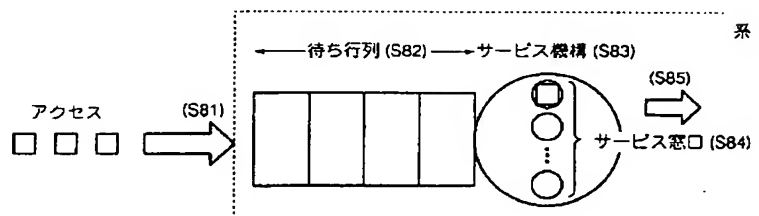
【図7】

予測モデル	係数	相関係数
<線形モデル> $y = ax + b$	$a = 0.566$ $b = 0.822$	0.96
<二乗モデル> $y = ax^2 + b$	$a = 0.052$ $b = 1.936$	0.99
<ルートモデル> $y = a\sqrt{x} + b$	$a = -1.240$ $b = 2.304$	0.91

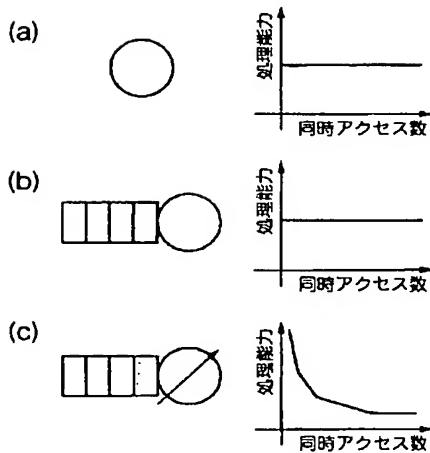
【図6】



【図8】

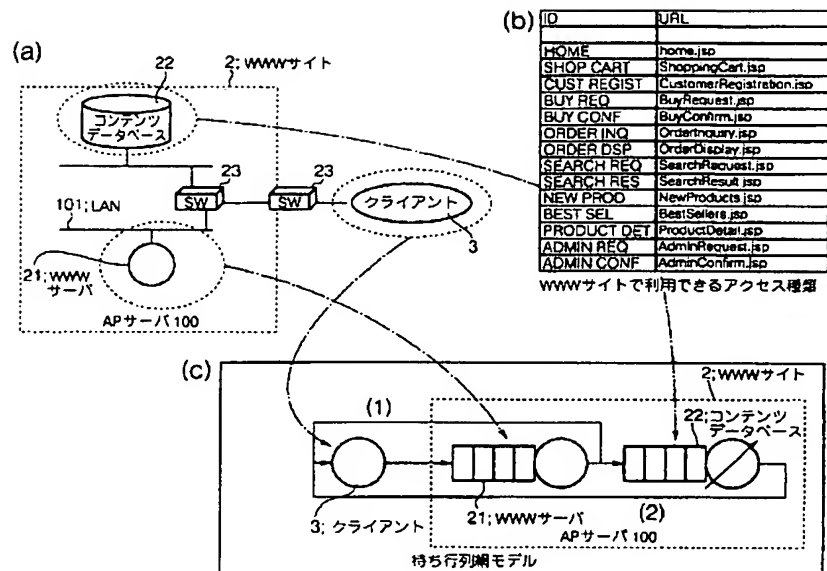


【図 9】



【図 11】

【図 10】



【図 12】

URL	動作タイプ	負荷 (%)
home.jsp	(2)	WWW/APサーバ 0.03 コンテンツDB 0.75
ShoppingCart.jsp	(2)	WWW/APサーバ 0.03 コンテンツDB 1.26
CustomerRegistration.jsp	(1)	WWW/APサーバ 0.03 コンテンツDB -
BuyRequest.jsp	(2)	WWW/APサーバ 0.03 コンテンツDB 0.65
BuyConfirm.jsp	(2)	WWW/APサーバ 0.03 コンテンツDB 1.57
OrderInquiry.jsp	(1)	WWW/APサーバ 0.03 コンテンツDB -
OrderDisplay.jsp	(2)	WWW/APサーバ 0.03 コンテンツDB 1.18
SearchRequest.jsp	(2)	WWW/APサーバ 0.03 コンテンツDB 0.54
SearchResult.jsp	(2)	WWW/APサーバ 0.03 コンテンツDB 1.74
NewProducts.jsp	(2)	WWW/APサーバ 0.03 コンテンツDB 2.83
BestSellers.jsp	(2)	WWW/APサーバ 0.03 コンテンツDB 4.78
ProductDetail.jsp	(2)	WWW/APサーバ 0.03 コンテンツDB 0.33
AdminRequest.jsp	(2)	WWW/APサーバ 0.03 コンテンツDB 0.31
AdminConfirm.jsp	(2)	WWW/APサーバ 0.03 コンテンツDB 0.66

WWW 21/APサーバ 100	1.00
コンテンツDB 22	8.00

(c)	最大待ち行列数	即時処理サービス数	共有リソース	待ち時間
WWWサーバ 21/APサーバ 100	最大TCPコネクション数	サーバの最大同時スレッド数		
コンテンツDB 22		WWWサーバ/APサーバのデータベースに対する最大コネクション数		
クライアント 3			1	TPC-Wの仕様に従った 間待ち時間

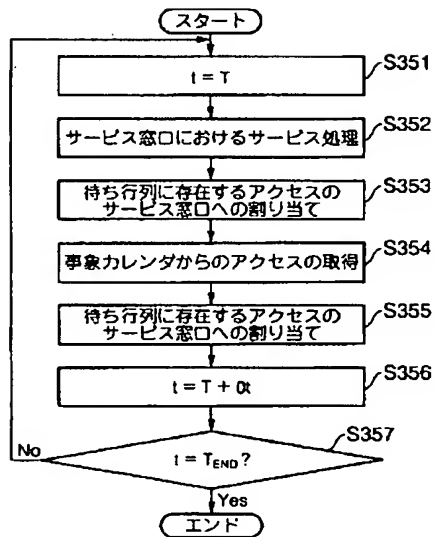
【図 14】

事象カレンダー 120

時間	アクセス
2/15 10:00:01	HOME
2/15 10:00:01	SHOP CART
2/15 10:00:01	CUST REGIST
2/15 10:00:02	HOME
2/15 10:00:02	BUY CONF
2/15 10:00:02	ORDER INQ
2/15 10:00:03	HOME
2/15 10:00:03	SEARCH REQ
2/15 10:00:03	SEARCH RES
2/15 10:00:03	NEW PROD
2/15 10:00:03	BEST SEL
2/15 10:00:03	HOME
2/15 10:00:04	SEARCH REQ
2/15 10:00:04	SEARCH REQ
...	...

ID	待ち行列における系内時間		レスポンスタイム
	WWW/APサーバ	コンテンツDB	
HOME	0.05	7.6	7.65
SHOP CART	0.05	12.7	12.75
CUST REGIST	0.05	0	0.05
BUY REQ	0.05	6.6	6.65
BUY CONF	0.05	15.8	15.85
ORDER INQ	0.05	0	0.05
ORDER DSP	0.05	11.9	11.95
SEARCH REQ	0.05	5.5	5.55
SEARCH RES	0.05	17.5	17.55
NEW PROD	0.05	28.4	28.45
BEST SEL	0.05	47.9	47.95
PRODUCT DET	0.05	3.4	3.45
ADMIN REQ	0.05	3.2	3.25
ADMIN CONF	0.05	6.7	6.75

【図 13】



フロントページの続き

F ターム(参考) 5B042 GA12 HH07 HH20 MA08 MA14
 MB03 MC22 MC25 MC28 MC35
 5B089 GA01 GB02 HA10 JB16 KA06
 KA07 KA12 KB04 KC48 MC02